DNA SEQUENCES IN A CLONED RESTRICTION FRAGMENT CONTAINING THE INVERTED REPEAT REGION FROM YEAST 2 μm PLASMID

J. HINDLEY

Department of Biochemistry, University of Bristol, University Walk, Bristol BS8 1TD, England

Received 10 August 1977

1. Introduction

While the presence of plasmid DNA is widespread in bacteria it is uncommon among eukaryotes. Recently a novel class of hereditary factors responsible for drug resistance in yeast, *Saccharomyces cerevisiae,* have been demonstrated. By some criteria (2:2 meiotic segregation, centromere linkage) they behave like nuclear genes while by other criteria (loss during vegetative multiplication, apparent migration from one chromosome to another) they behave like bacterial plasmids or episomes [1]. One of these factors, which carries an oligomycin resistance determinant, shows a direct correlation with the presence of a 2.2 μm covalently closed circular DNA molecule. There are 50–80 such molecules per haploid genome and these constitute about 3% of the total DNA. Guerineau et al. [2] have further shown that the 2.2 μm plasmid contains two 600 base-pair inverted repeat sequences separated by non-repeated DNA. This presence of inverted repeat sequences and the correlation of the presence of the plasmid with oligomycin resistance makes analogies to bacterial-transposable antibiotic-resistance determinants an attractive possibility. Inverted repeat sequences in bacterial plasmids have been shown to promote the insertion and/or translocation of defined genetic units at several sites on either the plasmid or the bacterial chromosome [3–6]. Inverted repeats have also been demonstrated in bacterial [7] and animal viruses [8]; the presence of such inverted repeats in the yeast plasmid may therefore be responsible for its unusual genetic properties.

The 2.2 μm DNA as isolated from yeast is a mixture of two types of molecule resulting from recombination between the inverted repeats. This has been clearly demonstrated, e.g., by the observation that the molecules have two *Eco*RI targets asymmetrically disposed with respect to the inverted repeats such that digestion of the mixture of these two forms with *Eco*RI results in four fragments of different size [2]. Beggs et al. [9] and Hollenberg et al. [10] have further mapped the targets for *Eco*RI, *Hin*dIII and *Hpa*I on both forms of the plasmid. Several of the *Hin*dIII fragments have now been cloned in phage λ [10,11]. One of these recombinant phages contains an inserted 0.6×10^6 dalton fragment (approx. 900 base pairs) corresponding to the whole of the 600 base-pair inverted repeats together with a flanking region of about 300 base pairs extending into the 'S' loop of the non-repeated DNA [9]. This recombinant therefore provides a convenient source of DNA for sequence analysis of the inverted repeat region. This λ clone, kindly provided by Dr John F. Atkins, was used in the experiments described.

2. Experimental

The recombinant phage was grown in 1 litre cultures of *E. coli* strain QR 48 in L broth supplemented with 10^{-2} M magnesium sulphate. Following lysis the phage was precipitated by the addition of polyethylene glycol 6000 (10%) and NaCl to 0.5 M and purified by banding in 42% (w/w) caesium chloride. After dialysis phage DNA was extracted with phenol and dialysed extensively against 10 mM Tris (pH 8.0), 1 mM EDTA. Yields of 3–4 mg phage DNA/ litre lysate were obtained. Digestion of the DNA with *Hin*dIII was

used to excise the cloned segment which was subsequently separated from the right and left arms of λ by preparative gel electrophoresis. For sequence analysis the *Hin*dIII fragment was eluted from gels, in the absence of any carrier, and incubated with alkaline phosphatase to generate free 5′-hydroxyl ends. These were subsequently labelled using [γ-^{32}P]ATP and polynucleotide kinase. The fragment, now labelled at both ends, was cut into two pieces of chain lengths 150 and 750 base pairs, respectively, by incubation with the restriction enzyme *Hae*III. These were finally separated by electrophoresis in a 6% acrylamide gel. Following elution the sequence adjacent to the 5′-labelled ends were determined using the method of Maxam and Gilbert [12].

3. Results and discussion

Figure 1 shows the result of one sequencing gel from which the sequence of the first 30 nucleotides from the 5′-end of the smaller of the two fragments can be determined. In order to resolve the larger sized oligonucleotides further sequencing gels were run in which the bromophenol blue and xylene cyanol markers were successively electrophoresed to the bottom of the gel. This gave adequate resolution for determining sequences of 99 and 83 nucleotides respectively from the 5′-labelled ends of the 150 and 750 base pair fragments. Figure 2 shows the sequences obtained for the 5′-phosphorylated single strand with the complementary sequence written below. The 150 base pair fragment is seen to contain an *Hha*I cleavage site at position 36–37 and a *Taq*I site at position 44–45. Both these fragments have subsequently been isolated and their positions and sequences confirmed.

At present no conclusions can be drawn about the

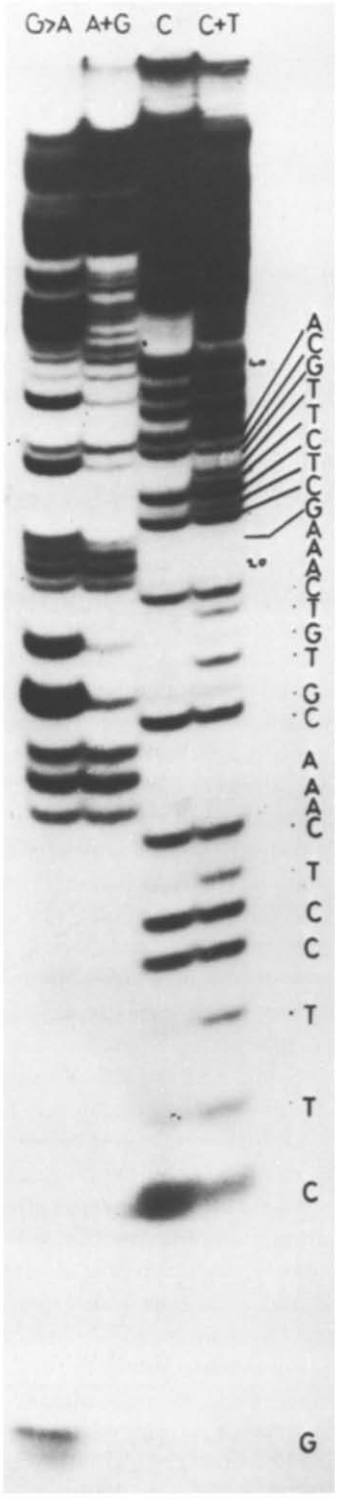

Fig.1. Autoradiogram of a sequencing gel showing the separation of [^{32}P]-end labelled oligonucleotides obtained by base specific cleavages of the smaller (150 base pair) fragment. Chemical degradation was performed as described by Maxam and Gilbert [12]. The sequence of the first 29 bases is shown at the right of the autoradiogram. The first nucleotide of the labelled 5′-end is destroyed by the base specific cleavage and is not present on the autoradiogram. Its nature was deduced from the known specificity of *Hin*dIII.

biological function of these sequences since nothing is known about their coding or transcriptional potential in the plasmid. However some interesting features are apparent. In the smaller fragment I there is considerable sequence duplication. Thus the sequence TCAAA found in position 8–12 is repeated in position 17–21. The hexanucleotide TCTTGC occupies positions 24–29 and 33–38 and the heptanucleotide TTGTTAT is found in 49–55 and again in 69–75. Another heptanucleotide TGTACGC is found in positions 55–61 and 86–92. The 83 nucleotide sequence II from the larger fragment does not show these features but two nine base palindromic sequences are found, namely GACGCGCAG in position 31–39 and AGTTATTGA in position 65–73. Whether these represent binding sites for regulatory proteins is not known. The experiments so far do not reveal the position of the inverted repeat segment within the cloned DNA. Preliminary data from a further λ-cloned sequence suggests the 150 base fragments contains the left hand

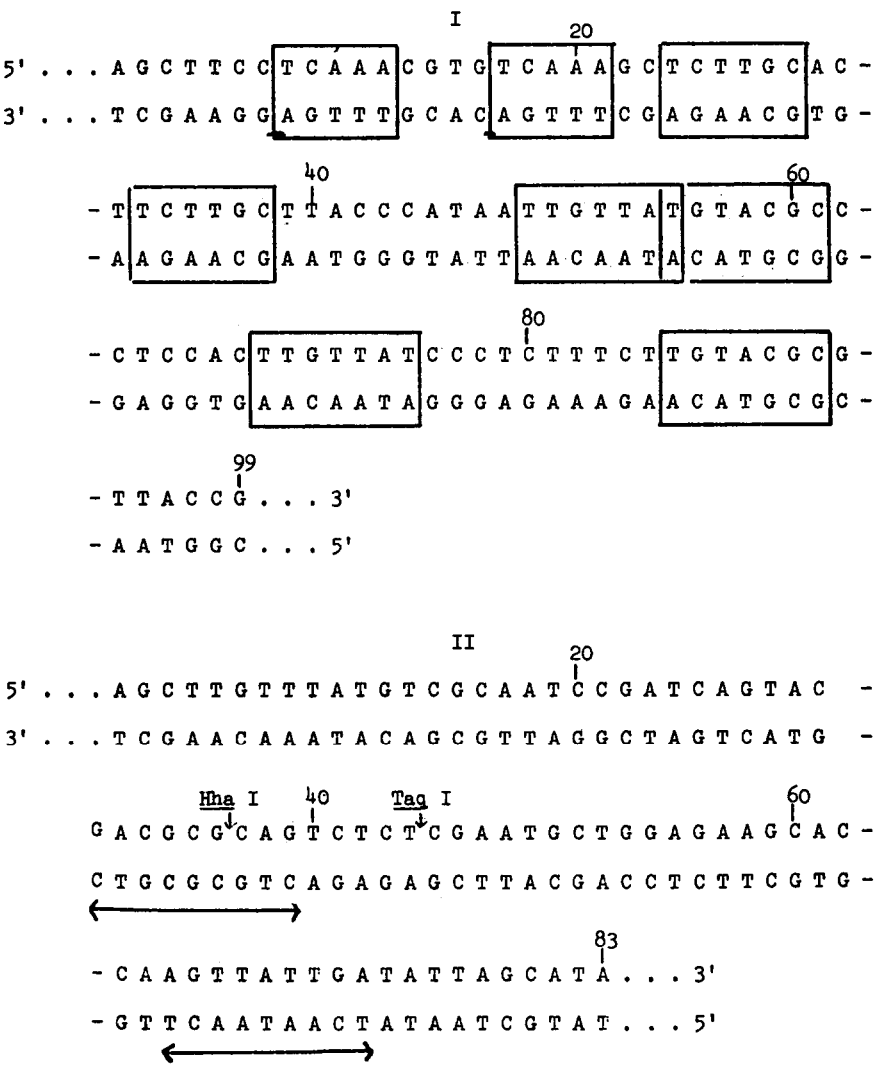

Fig.2. Nucleotide sequences adjacent to the 5'-ends of the cloned *Hin*dIII fragment derived from yeast 2.2 μm plasmid. I is the 5'-sequence from the smaller and II from the larger of the *Hae*III fragments. Boxed-in regions show repeated sequences and palindromic sequences are indicated by arrows.

end of the inverted repeat and the longer fragment arises from the non-repeated flanking region. Further experiments are in progress to confirm this and define precisely the ends of the inverted repeat segment. Further sequence analysis, currently in progress, is aimed at determining the entire 900 base-pair sequence. At present over 300 nucleotides have been sequenced using the 5′-ends exposed by further restriction enzymes. However the data for linking these together is at present incomplete. Further experiments involving RNA polymerase binding and in vitro transcription will be necessary to identify the function of these sequences within the plasmid.

Acknowledgements

This research was supported by a grant from the Science Research Council. I thank Dr J. F. Atkins, Department of Molecular Biology, University of Edinburgh for the phage λ clone used in this work, Drs A. Maxam and W. Gilbert for providing a description of their DNA-sequencing protocol prior to publication, and G. A. Phear for skilled assistance.

References

[1] Guerineau, M., Slonimski, P. P. and Avner, P. (1974) Biochem. Biophys. Res. Commun. 61, 462–469.
[2] Guerineau, M., Grandchamp, C. and Slonimski, P. P. (1976) Proc. Natl. Acad. Sci. USA 73, 3030–3034.
[3] Sharp, P. A., Cohen, S. N. and Davidson, N. (1973) J. Mol. Biol. 75, 235–255.
[4] Saedler, H., Reif, H. J., Hu, S. and Davidson, N. (1974) Mol. Gen. Genet. 132, 265–289.
[5] Kopecko, D. J. and Cohen, S. N. (1975) Proc. Natl. Acad. Sci. USA 72, 1373–1377.
[6] Hoffman, F., Rubens, C. and Falkow, S. (1975) Proc. Natl. Acad. Sci. USA 72, 3623–3627.
[7] Hsu, M. and Davidson, N. (1974) Virology 58, 229–239.
[8] Sheldnick, P. and Berthelot, N. (1974) Cold Spring Harbor Symp. Quant. Biol. 39, 667–678.
[9] Beggs, J. D., Guerineau, M. and Atkins, J. F. (1976) Mol. Gen. Genet. 148, 287–294.
[10] Hollenberg, C. P., Degelmann, A., Kustermann-Kuhn, B. and Roger, H. D. (1976) Proc. Natl. Acad. Sci. USA 73, 2072–2076.
[11] Atkins, J. F. (1976) Personal communication.
[12] Maxam, A. and Gilbert, W. (1977) Proc. Natl. Acad. Sci. USA 74, 560–564.